STAT C206B: Topics in Stochastic Processes

Feynman Liang*
Department of Statistics, UC Berkeley

Last updated: January 23, 2020

Contents

1 Lecture 1: Backround material

2020-01-21

1.1 Ferguson distributions / Dirichlet processes

Definition 1 (*Gamma distribution*)

Random variable X supported on $(0, \infty)$ has ***Gamma distribution*** with shape $\alpha > 0$ and inverse scale / rate $\beta > 0$, written $X \sim \text{Gamma}(\alpha, \beta)$ if it has density

$$f_X(t) = \mathbb{1}\{t \in (0, \infty)\} \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \tag{1}$$

where $\Gamma(t) = \int_0^\infty u^{t-1} e^{-u} du$ is the Gamma function defined for all $\Re t > 0$ and analytically continued to $\mathbb{C} \setminus \{n \in \mathbb{Z} : n < 0\}$

Proposition 2 (*Gamma closed under summation*)

{prop:gamma-closed-sum} *If $Y \sim \text{Gamma}(\alpha, \beta)$ and $Z \sim \text{Gamma}(\gamma, \beta)$ are independent, then $Y + Z \sim \Gamma(\alpha + \gamma, \beta)$.*

Proof.

$$\begin{aligned}
f_{Y+Z}(t) &= \int_0^t f_Y(u) f_Z(t-u) du \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\gamma)} \beta^{\alpha+\gamma} e^{-\beta t} \int_0^t u^{\alpha-1} (t-u)^{\gamma-1} du \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\gamma)} \beta^{\alpha+\gamma} e^{-\beta t} \int_0^1 (tv)^{\alpha-1} (t-(tv))^{\gamma-1} t dv \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\gamma)} \beta^{\alpha+\gamma} e^{-\beta t} t^{\alpha+\gamma-1} B(\alpha, \gamma)
\end{aligned}$$

where $B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$ is the beta function $\square$

A closely related distribution obtained from concatenating Gamma random variables into a vector and then normalizing the sum to 1 is the Dirichlet distribution.

*feynman@berkeley.edu

Definition 3 (*Dirichlet distribution*)

Let $\boldsymbol{\alpha} \in (0, \infty)^K$. Random (probability) vector X taking values on the $K-1$-dimensional probability simplex $\Delta^{K-1} = \{\boldsymbol{x} \in [0,1]^K : \sum_i x_i = 1\}$ has ***Dirichlet distribution*** of order K and concentration parameters $\boldsymbol{\alpha}$, denoted $X \sim \mathrm{Dir}(\boldsymbol{\alpha})$, if it has density

$$f_X(\boldsymbol{x}) = \mathbb{1}\{\boldsymbol{x} \in \Delta\} \underbrace{\frac{\Gamma\left(\sum_{i=1}^K \alpha_i\right)}{\prod_{i=1}^K \Gamma(\alpha_i)}}_{=:B(\boldsymbol{\alpha})^{-1}} \prod_{i=1}^K x_i^{\alpha_i - 1}$$

Proposition 4 (*Constructing Dirichlet from Gammas*)

{prop:dirichlet-from-gamma} *Let $X_1, \ldots, X_n$ be independent* Gamma(α_i, β) *distributed, $S_n = \sum_{i=1}^n X_i$. Then $(V_i)_i = (X_i/S_n)_i \sim$* $\mathrm{Dir}(\boldsymbol{\alpha})$.

Proof. $S_n \sim \Gamma(\sum_i^n \alpha_i, \beta)$ by Proposition 2 and for $\boldsymbol{v} \in \Delta^{n-1}$, we have

$$\begin{aligned} f_V(\boldsymbol{v}) &= \int_0^\infty f_X\left(sv_1, \ldots, sv_{n-1}, sv_n\right) f_{S_n}(s) ds \\ &= \int_0^\infty e^{-\sum_{i=1}^n sv_i} \left(\prod_{i=1}^n \frac{(sv_i)^{\alpha_i - 1}}{\Gamma(\alpha_i)} \right) \frac{s^{\sum_i^n \alpha_i - 1} e^{-s}}{\Gamma(\sum_i^n \alpha_i)} ds \\ &= \frac{1}{\prod_1^n \Gamma(\alpha_i)} \prod_{i=1}^n v_i^{\alpha_i - 1} \int_0^\infty e^{-s\cancelto{1}{\sum_1^n v_i}} s^{(\sum_1^n \alpha_i) - 1} ds \\ &= \frac{\Gamma(\sum_{i=1}^n \alpha_i)}{\prod_1^n \Gamma(\alpha_i)} \prod_{i=1}^n v_i^{\alpha_i - 1} \end{aligned}$$

□

Similar to Proposition 2 (Gamma closed under summation), where adding two Gammas yielded another Gamma where the parameters were added, Dirichlet distributions enjoy a similar kind of closure: "clumping" coordinate axes together (described below) yields another Dirichlet distribution where the parameters of the clumped axes are summed together.

Proposition 5 (*Dirichlet clumping property*)

Suppose $X \sim$ $\mathrm{Dir}(\alpha_1, \ldots, \alpha_n)$. *For any $r \leq n$, let $V_i = X_i$ for $i \in [r]$ and let $V_{r+1} = \sum_{j=r+1}^n X_j$. Then $V \sim$* $\mathrm{Dir}(\alpha_1, \ldots, \alpha_r, \sum_{j=r+1}^n \alpha_j)$.

Proof. By induction, it suffices to show this for $r = n-2$. Notice

$$\begin{aligned} f(v_1, \ldots, v_r, s) &= B(\boldsymbol{\alpha})^{-1} \left(\prod_{i=1}^{n-1} v_i^{\alpha_i - 1} \right) \int \mathbb{1}\left\{x_{n-1} + x_n = s\right\} x_{n-1}^{\alpha_{n-1} - 1} x_n^{\alpha_n - 1} dx_{n-1} dx_n \\ &= B(\boldsymbol{\alpha})^{-1} \left(\prod_{i=1}^{n-1} v_i^{\alpha_i - 1} \right) \int_0^s u^{\alpha_{n-1} - 1} (s-u)^{\alpha_n - 1} du \\ &= B(\boldsymbol{\alpha})^{-1} \left(\prod_{i=1}^{n-1} v_i^{\alpha_i - 1} \right) s^{\alpha_{n-1} + \alpha_n - 1} B(\alpha_{n-1}, \alpha_n) \end{aligned}$$

Since $\frac{B(\alpha_{n-1}, \alpha_n)}{B(\boldsymbol{\alpha})} = \frac{\Gamma(\sum_1^n \alpha_i)}{\Gamma(\alpha_{n-1} + \alpha_n) \prod_1^{n-2} \Gamma(\alpha_i)}$, we are done. □

Iterating this result over coordinate axes other than the last $n-r$, we see that "clumping together" entries in a Dirichlet random vector yields another Dirichlet random vector with parameters also "clumped together." Concretely, for any mapping $\phi : [n+1] \to [m+1]$ if $U_j = \sum_{\phi(i)=j} V_i$ then U has Dirichlet distribution with parameters $\gamma_j = \sum_{\phi(i)=j} \alpha_i$.

Generalizing this clumping property is the motivation for ***Ferguson Distributions*** ferguson1973 (Ferguson, 1973).

Definition 6 (*Ferguson / Dirichlet process distribution*)

Let μ be a finite positive Borel measure on complete separable metric space E. A random probability measure μ^* on E (i.e. a stochastic process indexed by a σ-algebra on E) has ***Ferguson distribution with parameter*** μ if for every finite partition $(B_i)_{i\in[r]}$ of E the random vector

$$(\mu^*(B_i))_{i\in[r]} \sim \mathrm{Dir}(\mu(B_1),\dots,\mu(B_r))$$

Lemma 7 (*Preservation of Ferguson under pushforward*)

Let μ^* be Ferguson with parameter μ and $\phi : E \to F$ measurable. Then the pushforward $\mu^* \circ \phi^{-1}$ is a random probability measure on F that has Ferguson distribution with parameter $\mu \circ \phi^{-1}$.

Proof. For $(B_i)_{i\in[r]}$ a finite partition of F, $(\phi^{-1}(B_i))_i$ is a finite partition of E. Since μ^* is Ferguson

$$(\mu^*(\phi^{-1}(B_i)))_i \sim \mathrm{Dir}((\mu(\phi^{-1}(B_i)))_i)$$

Hence $\mu^* \circ \phi^{-1}$ is Ferguson with parameter $\mu \circ \phi^{-1}$. □

Next, we turn to an important class of a Ferguson distributions arising from generalizing the Pólya urn.

Definition 8 (*Polya sequence*)

A sequence $(X_n)_{n\in\mathbb{N}}$ with values in E is a ***Polya sequence with parameter*** μ if for all $B \subset E$.

$$\Pr[X_1 \in B] = \mu(B)/\mu(E)$$
$$\Pr[X_{n+1} \in B \mid X_1,\dots,X_n] = \mu_n(B)/\mu_n(E)$$

where $\mu_n = \mu + \sum_{i=1}^n \delta_{X_i}$.

Remark 9. When E is finite (e.g. a set of colors for the balls), (X_n) represents the result of successive draws from an urn with initially $\mu(x)$ balls of color $x \in E$ and after each draw a ball of the same color as the one drawn is added back to give an urn with color distribution $\mu_{n+1}(x)$.

blackwell1973
Blackwell and MacQueen (1973) gives the following result connecting Pólya sequences and Ferguson distributions.

Theorem 10 (*Polya Urn Schemes*)

Let (X_n) be a Polya sequence with parameter μ. Then:

1. *$m_n = \mu_n/\mu_n(E)$ converges almost surely to a limiting discrete measure μ^**
2. *μ^* has Ferguson distribution with parameter μ*
3. *Given μ^*, $(X_i)_{i\geq 1}$ are independent with distribution μ^**

Proof. First consider E finite and let μ^* and $\{X_i\}$ be random variables whose joint distribution satisfies (2.) and (3.). Let π_n be empirical distribution of $(X_i)_{i\in[n]}$. $X_i \overset{\text{iid}}{\sim} \mu^*$, so by SLLN $\pi_n \overset{as}{\to} \mu^*$ and since

$$m_n = \frac{\mu + n\pi_n}{\mu(E) + n} \tag{2}$$

(1.) follows.

It remains to show (X_n) is a Polya sequence with parameter μ, i.e.

$$\Pr[A] = \prod_x \mu(x)^{[n(x)]}/\mu(E)^{[n]} \tag{3}$$ {{eq:polya-seq-meas}}

where $A = \{X_i = x_i\}_i$, $n(x) = \#\{i : x_i = x\}$, and the rising factorial $a^{[k]} = a(a+1)\cdots(a+k-1)$. Notice

$$\Pr[A] = \mathbb{E}\left[\Pr[A \mid \mu^*]\right] = \mathbb{E}\left[\prod_x \mu^*(x)^{n(x)}\right] \tag{4}$$

Since μ^* is Ferguson, viewing $E = \sqcup_{x\in E}\{x\}$ as a partition we have $(\mu^*(x))_{x\in E} \sim \mathrm{Dir}((\mu(x))_{x\in E})$ so the RHS is the $(n(x))_{x\in E}$ moment of the Dirichlet distribution, which is equal to

$$\mathbb{E}\left[\prod_x \mu^*(x)^{n(x)}\right] = \frac{\Gamma(\mu(E))}{\Gamma(\mu(E)+n)}\prod_x \frac{\Gamma(\mu(x)+n(x))}{\Gamma(\mu(x))} = \frac{1}{\mu(E)^{[n]}}\prod_x \mu(x)^{[n(x)]} \tag{5}$$

as required by Eq. (3).

General E follows from approximation argument. □

We leave the discreteness part of (1.) as an exercise, noting that similar to how Dirichlets can be defined as a set of independent Gammas normalized by their sum (Proposition 4 (Constructing Dirichlet from Gammas)) we would expect the Dirichlet process / Ferguson random measures to be definable as a gamma process with independent "increments" divided by their sum.

Exercise 11. Prove every Ferguson random measure is discrete. (Hint: argue using moments).

Remark 12. If (X_n) a Polya sequence, then it is a mixture of iid sequences, and is exchangeable i.e. $(X_i) \overset{d}{=} (X_{\sigma(i)})$ (see Eq. (3))

1.2 Construction of Haar Measure

For a finite group G, the measure $\mu(g) = \frac{1}{\#G}$ is left and right translation invariant i.e. $\mu(gA) = \mu(A) = \mu(Ag)$.

In fact, all compact groups have unique translation invariant measure called the Haar measure. For example $Z_{ij} \overset{\text{iid}}{\sim} N(0,1)$ for $i,j \in [n]$ and X the Gram-Schmidt orthonormalizataino of the rows of Z. Then $XU \overset{d}{=} UX$ for all $U \in O(n)$, so X has Haar measure on compact group $O(n)$.

Definition 13

A ***topological vector space*** (TVS) is a vector space equipped with a topology such that vector space operations are jointly continuous.

Example 14

$\mathbb{R}^n$ with standard topology, any Banach space.

Definition 15

A family $\mathfrak{G}$ of linear transformations on TVS $\mathfrak{X}$ is ***equicontinuous on subset*** $K \subset \mathfrak{X}$ if for every neighborhood V of the origin, there exists a neighborhood U of the origin such that

$$\forall k_1, k_2 \in K : k_1 - k_2 \in U \Rightarrow \mathfrak{G}(k_1 - k_2) \subset V \tag{6}$$

That is, $T(k_1 - k_2) \in V$ for all $T \in \mathfrak{G}$.

Definition 16

A ***locally convex topological vector space*** (LCTVS) is a TVS with a local base of absolutely convex absorbing sets at the origin.

To construct Haar measure for any compact group, we will need a fix point theorem due to Kakutani.

Theorem 17 (*Kakutani Fix Point Theorem*)

{thm:kakutani} *K compact convex subset of LCTVS $\mathfrak{X}$, $\mathfrak{G}$ group of linear transforms equicontinuous on K and such that $\mathfrak{G}(K) \subset K$, then $\exists p \in K$ such that*

$$\mathfrak{G}(p) = \{p\} \tag{7}$$

Proof.
- By Zorn's lemma applied to chains $(K_\lambda)_\lambda$ (note $K_a \subset K_b$ for $a < b$), $\exists$ minimal $K_1 \subset K$ such that $K_1 \neq \emptyset$ and $\mathfrak{G}(K_1) \subset K_1$.

- If K_1 is a single point, then proof is complete.
- Otherwise, by minimality the compact (becacuse $-$ is continuous) set $K_1 - K_1$ contains a point other than the origin, so exists $V \in N(0)$ such that $\bar{V} \not\supset K_1 - K_1$.
- For some $|\alpha| \leq 1$, there is a convex neighborhood $V_1 \in N(0)$ such that $\alpha V_1 \subset V$.
- By equicontinuity of $\mathfrak{G}$ on $K \supset K_1$, there is $U_1 \in N(0)$ such that for $k_1, k_2 \in K$ and $k_1 - k_2 \in U_1$ we have $\mathfrak{G}(k_1 - k_2) \subset V_1 >$
- Because $T \in \mathfrak{G}$ is invertible ($\mathfrak{G}$ is a group), T maps open sets to open sets and $T(A \cap B) = TA \cap TB$ for sets A, B.
- Since T is linear, for any A

$$T \operatorname{conv}(A) = \operatorname{conv}(TA) \tag{8}$$

- So

$$U_2 := \operatorname{conv}(\mathfrak{G}U_1 \cap (K_1 - K_1)) = \operatorname{conv}(\mathfrak{G}(U_1 \cap (K_1 - K_1))) \subset V$$

  is relatively open in $K_1 - K_1$ and satisfies $\mathfrak{G}U_2 = U_2 \not\supset K_1 - K_1$.
- By continuity, $\mathfrak{G}U_2 = \overline{\mathfrak{G}U_2}$.
- Let $\delta = \inf\{a : a > 0, aU_2 \supset K - 1 - K_1\} \geq 1$, by compactness $\delta < \infty$. Let $U := \delta U_2$.
- For each $\varepsilon \in (0, 1)$

$$(1 + \varepsilon)U \supset K_1 - K_1 \not\subset (1 - \varepsilon)\overline{U}$$

- Because $(1 - 1/4n)\bar{U} \not\supset K_1 - K_1$, we have $K_2 \neq K - 1$.
- K_2 is closed and convex
- Further, since $T(a\bar{U}) \subset a\bar{U}$ for $T \in \mathfrak{G}$, we have

$$T(a\bar{U} + k) \subset a\bar{U} + Tk \qquad \text{forall } T \in \mathfrak{G}, k \in K_1$$

- Recalling $TK_1 \subset K_1$ for $T \in \mathfrak{G}$ and that $\mathfrak{G}$ is a group, we find that $TK_1 = K_1 \mathfrak{G} K_2 =$ a contradiction Finish

□

Theorem 18 (***Existence of Haar Measure***)

{thm:haar-measure}

G compact group, $\mathcal{C}(G)$ continuous maps $G \to \mathbb{R}$. Then there is a unique linear form $m : \mathcal{C}(G) \to \mathbb{R}$ such that

1. *$m(f) \geq 0$ for $f \geq 0$ (positive)*
2. *$m(\mathbb{1}) = 1$ (normalized)*
3. *$m({}_sf) = m(f)$ where ${}_sf(g) = f(s^{-1}g)$ for $s, g \in G$ (left invariant)*
4. *$m(f_s) = m(f)$ where $f_s(g) = f(gs)$ (right invariant)*

2 Lecture 2: Measure theory continued

2020-01-23

Today we will start off by completing the construction of the Haar measure. All metric spaces are Hausdorff.

Recall that Theorem 17 (Kakutani Fix Point Theorem) requires equicontinuity of a set of continuous maps. The following result gives an alternative characterization of equicontinuity.

Theorem 19 (***Generalization of Arzela-Ascoli***)

{thm:arzela-ascoli}

Let X be a compact topological space. A subset of function $F \subset \mathcal{C}(X)$ is relatively compact in topology induced by uniform norm $\Leftrightarrow$ F is equicontinuous and pointwise bounded.

$f \mapsto {}_sf$ linear transform of $\mathcal{C}(G)$, $\|{}_sf - {}_sg\| = \|f - g\|$, G acts as a group of isometries on $\mathcal{C}(G)$. In particular, this group is equicontinuous.

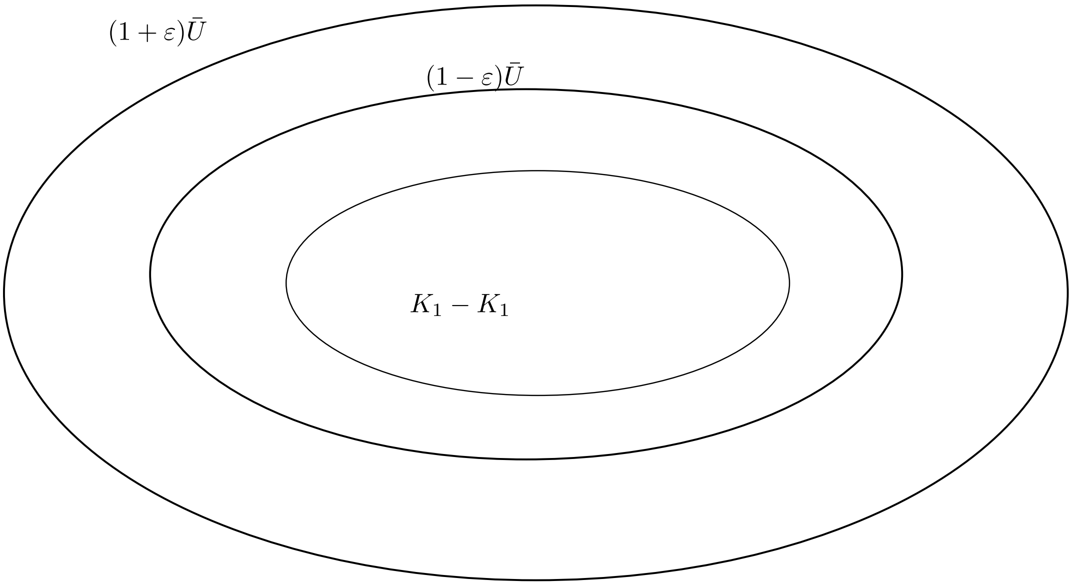


Figure 1: Sketch of proof of Kakutani's theorem

Proof of Theorem 18. Fix $f \in \mathcal{C}(G)$. Let $\mathcal{C}_f$ denote the convex hull of all left translates of f. Elements $g \in \mathcal{C}_f$ are finite sums of form

$$g(x) = \sum_{\text{finite}} a_i f(s_i x), \qquad a_i > 0, \sum_{\text{finite}} a_i = 1$$

Clearly $\|g\|_\infty = \max\{|g(x)| : x \in G\} \leq \|f\|_\infty$, thus sets of the form $\mathcal{C}_f(x) = \{g(x) : g \in \mathcal{C}_f\}$ are bounded.

Since G is compact, f is uniformly continuous hence for $\varepsilon > 0$ there exists a neighborhood $V = V_\varepsilon$ of the identity $e \in G$ such that

$$y^{-1}x \in V \Rightarrow |f(x) - f(y)| \leq \varepsilon$$

Since $(s^{-1}y)^{-1}s^{-1}x = y^{-1}x$, we also have .

By Theorem 19 (Generalization of Arzela-Ascoli), $\mathcal{C}_f$ relatively compact in $\mathcal{C}(G)$, so its closure $K_f = \overline{\mathcal{C}_f}$ is compact convex. G acts by left translation on $\mathcal{C}(G)$ and leaves $\mathcal{C}_f$ invariant, hence K_f invariant as well. By Theorem 17 (Kakutani Fix Point Theorem), there is a fixed point $g \in K_f$ of this action of G on K_f which satisfies

$$_sg = g \ (\forall s \in G) \quad \Rightarrow \quad g(s^{-1}) = {_sg}(e) = g(e) = c \ (\forall s \in G)$$

for some constant c.

By definition of K_f, for any $\varepsilon > 0$ there is $\{s_1, \ldots, s_n\} \subset G$ and $a_i > 0$ such that

$$\sum_{i=1}^n a_i = 1, \quad \text{and} \quad \left| c - \sum_{i=1}^n a_i f(s_i x) \right| < \varepsilon \qquad (\forall x \in G)$$

We first show there is only one constant function in K_f, so the fix point $g \in K_f$ is unique.

Start with same construction as before expect now use right translations of f (i.e. using the opposite group G' of G, or the function $f' = f(x^{-1})$, obtaining relatively compact set $\mathcal{C}'_f$ with compact convex closure K'_f containing constant function c').

Opposite group

The opposite group g' of the group G is the group that coincides with G as a set but has group operation $(x, y) \mapsto y^{-1}x^{-1}$

It will be enough to show $c = c'$, since all constants $c \in K_f$ must be equal to one chosen constant $c' \in K'_f$ and conversely.

By construction, exists finite combination of right translates close to c' i.e.

$$|c' - \sum_j b_j f(xt_j)| < \varepsilon \qquad \text{(for some } t_j \in G,\ b_j > 0 \text{ with } \sum_j b_j = 1) \tag{9}$$ {{eq:combo-close-to-constant}}

Multiply by a_i and put $x = s_i$ to get

$$|c' a_i - \sum_j a_i b_j f(s_i t_j)| < \varepsilon a_i$$

Summing over i

$$|c' - \sum_{i,j} a_i b_j f(s_i t_j)| = |c' \sum_i a_i - \sum_{i,j} a_i b_j f(s_i t_j)| < \varepsilon \sum_i a_i = \varepsilon$$

Operating symmetrically on Eq. (9) (mult by b_j, $x = t_j$) finish

The following properties are obvious:

- $m(\mathbb{1}) = 1$ since $K_f = \{1\}$ for $f = 1$
- $m(f) \geq 0$ if $f \geq 0$
- $m(af) = am(f)$ for any $a \in \mathbb{R}$ (since $K_{af} = K_f$)
- $m({}_s f) = m(f) = m(f_s)$ (by uniqueness)

It remains to show that m is additive (hence linear). Take $f, g \in \mathcal{C}(G)$ and start with Eq. (9) with $c = m(f)$, further letting $h(x) = \sum_i a_i g(s_i x)$. Since $h \in \mathcal{C}_g$, we have $\mathcal{C}_h \subset \mathcal{C}_g$ hence $K_h \subset K_g$. But K_g contains only one constant so in fact $m(h) = m(g)$.

We can write

$$|m(h) - \sum b_j h(t_j x)| < \varepsilon$$

for finitely many $t_j \in G$ and $b_j > 0$, $\sum_j b_j = 1$.

Using this definition of h and $m(h) = m(g)$ gives

$$|m(g) - \sum_{i,j} a_i b_j g(s_i t_j x)| < \varepsilon$$

However, multiplying Eq. (9) by b_j and replacing x , adding (9) and (10) gives

$$|m(f) + m(g) - \sum_{i,j} a_i b_j (f+g)(s_i t_j x)| < 2\varepsilon$$

Thus $m(f) + m(g) \in K_{f+g}$, establishing additivity. Note that the only constant in K_{f+g} is . $\square$???

We now want to head towards some integration against probability measures.

Definition 20

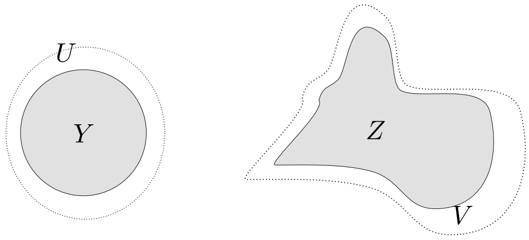
A topological space X is ***normal*** if for any disjoint closed sets Y and Z there exists disjoint open sets U and V such that $Y \subset U$ and $Z \subset V$.

Definition 21

X is ***completely regular*** if for all $y \in X$ and every closed $Z \subset X \setminus \{y\}$ there exists $f : X \to [0,1]$ continuous such that $f(y) = 0$ and $f(z) = 1$ for all $zinZ$.

Lemma 22 (***Urysohn***)

Every normal space is completely regular.



{fig:normal-topological-space} Figure 2: Normal topological spaces admit separating closed sets with two disjoint open sets

Definition 23

m a set function on X with values in $[0, +\infty]$ such that $m(\emptyset) = 0$. $A \subset X$ is ***Carathéodory measurable*** wrt m (Carathéodory m-measurable) if, for every $E \subset X$ we have

$$m(E \cap A) + m(E \setminus A) = m(E)$$

Let $\mathfrak{m}_m$ denote class of all Carathéodory m-measurable sets.

Theorem 24

1. *$\mathfrak{m}_m$ is an algebra, m is additive on $\mathfrak{m}_m$*
2. *For all sequences of pairwise disjoint $A_i \in \mathfrak{m}_m$: ???*
3. *If m is an outer measure on X, then $\mathfrak{m}_m$ is a σ-algebra, m is countably additive on $\mathfrak{m}_m$, and m is complete on $\mathfrak{m}_m$*

Example 25

Let $\mathfrak{X}$ be a family of subsets of X such that $\emptyset \in \mathfrak{X}$. Given $\tau : \mathfrak{X} \to [0, +\infty]$ with $\tau(\emptyset) = 0$, set

$$\mathfrak{m}(A) = \inf \left\{ \sum_{n=1}^{\infty} \tau(X_n) : X_n \in \mathfrak{X}, A \subset \cup_{n=1}^{\infty} X_n \right\}$$

where $\mathfrak{m}(A) = \infty$ in the absence of such sets X_n. Then $\mathfrak{m}$ is an outer measure, denoted τ^*.

The ***Borel σ-algebra***, $\mathcal{B}(X)$, is generated by all open sets.

Let $\mathcal{B}a(x)$ denote the ***Baire σ-algebra***, generated by

$$\{x \in X : f(x) > 0\} \tag{10}$$ {{eq:functionally-open}}

where f is a continuous function on X. It is the smallest σ-algebra where all continuous functions are measurable.

The sets of the form Eq. (10) are called ***functionally open***.

Lemma 26

In a metric space X, any closed set is the set of zeros of a continuous function. Hence, $\mathcal{B}(X) = \mathcal{B}a(X)$.

Lemma 27 (*Baire sets are countably determined*)

Every Baire set is determined by some countable family of functions, i.e. has the form

$$\{x : (f_i(x))_i : i < \infty, f_i \in \mathcal{C}(X)\}$$

A consequence of the monotone class theorem ???

Throughout, we consider (signed) measures of ***bounded variation*** unless explicitly denoted otherwise.

Definition 28

Let X be a topological space.

- A countably additive measure on $\mathcal{B}(X)$ is called a ***Borel measure***
- A countably additive measure on $\mathcal{B}a(X)$ is called a ***Baire measure***
- A Borel measure μ on X is called ***Radon measure*** if every $B \in \mathcal{B}(X)$ can be approximated from the inside by compact sets: for $\varepsilon > 0$ exists $K_\varepsilon \subset B$ such that $|\mu|(B \setminus K_\varepsilon) < \varepsilon$.

When are two Borel measures equal?

Lemma 29

If two Borel measures coincide on all open sets, thne they coincide on all Borel sets.

Proof. Split $\mu = \mu^+ - \mu^-$ and notice that each of the two components are nonnegative and coincide on open sets. By monotone class theorem, $\mu^+ = \mu^-$. □ finish

μ is Radon iff $|\mu|$ is Radon iff both μ^+ and μ^- are Radon.

Inner and outer approximatino of measures on $\mathbb{R}^n$:

Theorem 30

$\mu \geq 0$ on $\mathcal{B}(\mathbb{R}^n)$, then any Borel set $B \subset \mathbb{R}^n$ and any $\varepsilon > 0$ exists U_ε open and F_ε closed such that $F_\varepsilon \subset A \subset U_\varepsilon$ and $\mu(U_\varepsilon \setminus F_\varepsilon) < \varepsilon$.

Proof. Let $\mathcal{A}$ the class of all sets $A in \mathcal{B}$ such that $F_\varepsilon \subset A \subset U_e$ and $\mu(U_\varepsilon \setminus F_\varepsilon) < \varepsilon$ for some closed set F_ε and open set U_ε.

Every closed A is in $\mathcal{A}$, since we can take $F_\varepsilon = A$ and U_ε some open δ-nbd and consider $\delta \to 0$.

It suffices to show that $\mathcal{A}$ is a σ-algebra, since the closed sets generate $\mathcal{B}$. $\mathcal{A}$ is closed wrt complements, so it remains to verify closure under countable union.

Let $A_j \in \mathcal{A}$, $\varepsilon > 0$. Then exists closed F_j and open $U - j$ such that $F_j \subset A_j \subset U - j$ and $\mu(U_j \setminus F_j) < \varepsilon 2^{-j}$ for $j \in \mathbb{N}$.

The set $U = U_{j=1}^{\infty} U_j$ is open, and $Z_k = U_{j=1}^{k} F_j$ is closed.

Observe $Z_k \subset \cup_{j=1}^{\infty} A_j \subset U$ and for sufficiently large k $\mu(U \setminus Z_k) < \varepsilon$.

Indeed, $\mu(\cup_j^{\infty} \mu_j \setminus F_j) < 2$ □ ??

Definition 31

Set function $\mu \geq 0$ defined on $\mathcal{A} \subset 2^X$ is ***tight*** on $\mathcal{A}$ if $\forall \varepsilon > 0$ exists compact $K_\varepsilon \subset X$ such that $\mu(A) < \varepsilon$ for all $A \in \mathcal{A}$ that does not meet K_ε.
Additive set function μ of bounded variation on an algebra is ***tight*** if its total variation $|\mu|$ is tight.

A Borel measure is tight iff $\forall \varepsilon > 0$ exists compact K_ε such that $|\mu|(X \setminus K_\varepsilon) < \varepsilon$ (the "total variation measure").

The second definition is necessary to handle Baire sets.

Definition 32

μ is ***regular*** if $\forall A \in \mathcal{A}, \varepsilon > 0$, $\exists F_\varepsilon$ closed such that $F_\varepsilon \subset A$, $A \setminus F_\varepsilon \in \mathcal{A}$, and

Theorem 27 implies any Borel measure on $\mathbb{R}_n$ is regular, and the same proof works to show any Borel measure on metric space is regular.

Corollary 33 (*Baire measures are regular*)

Every Baire measure μ on topological space X is regular. Moreover, for every Baire set E and $\varepsilon > 0$, there exists a continuous function f on X such that $f^{-1}(0) \subset E$ and $|\mu|(E \setminus f^{1}(0)) < \varepsilon$.

Bibliography

{blackwell1973} Blackwell, D. and J. B. MacQueen
1973. Ferguson distributions via polya urn schemes. *Ann. Statist.*, 1(2):353–355.

{ferguson1973} Ferguson, T. S.
1973. A bayesian analysis of some nonparametric problems. *Ann. Statist.*, 1(2):209–230.